

# A lower bound on binomial tails: an approach via tail conditional expectations

Christos Pelekis\*

September 22, 2016

## Abstract

Using ideas from tail conditional expectations, we obtain a non-asymptotic lower bound on the probability that a binomial random variable is significantly larger than its expectation.

*Keywords:* lower bounds; binomial tails; tail conditional expectation

## 1 Prologue and main result

Given a positive integer  $n$  and a real number  $p \in (0, 1)$ , we denote by  $\text{Bin}(n, p)$  a binomial random variable of parameters  $n$  and  $p$ . Throughout the text, given two random variables  $X, Y$ , the notation  $X \sim Y$  will indicate that they have the same distribution. In this article we shall be interested in *lower bounds* on the “tail” probability  $\mathbb{P}[X \geq k]$ , where  $X \sim \text{Bin}(n, p)$  and  $k > np$ . The problem of estimating the tails of a binomial random variable is arguably one of the most basic problems in probability and statistics. It arises in a plethora of problems and, despite the fact that the binomial tails are asymptotically well understood (see, for example, [3, 4, 6, 12, 13, 14, 18, 20]), several questions coming from a variety of disciplines (see [1, 2, 8, 10, 16]) do not decrease the need for upper and lower bounds on  $\mathbb{P}[\text{Bin}(n, p) \geq k]$  that are valid for fixed  $n, p, k$ . Sharp upper bounds on binomial tails are provided by the, so-called, Chernoff-Hoeffding bound (see [9]) by employing an “exponential moment” approach:

$$\mathbb{P}[\text{Bin}(n, p) \geq k] \leq \exp(-nD(k/n||p)), \text{ for } k > np, \quad (1)$$

---

\*Charilaou Trikoupi 77, 30132, Agrinio, Greece; email: pelekis.chr@gmail.com

where  $D(k/n||p)$  is the Kullback-Leibler distance between  $k/n$  and  $p$ . Upper bounds that are sharper than the previous one can be obtained by employing a "factorial moment" approach (see [17]):

$$\mathbb{P}[\text{Bin}(n, p) \geq k] \leq p^\ell \binom{n}{\ell} / \binom{k}{\ell}, \text{ where } \ell = \lceil \frac{k - np}{1 - p} \rceil, k > np. \quad (2)$$

Here and later,  $\lceil x \rceil$  denotes the minimum integer that is larger than or equal to  $x$ . Elementary, though quite tedious, calculations (see [17] for details) show that the bound given by (2) is smaller than the bound given by (1) but the two bounds are quite close to each other.

A well-known theorem of Cramér implies that the bound of Eq.(1) is asymptotically tight and therefore one can only hope for sharp lower bounds on the binomial tails that are a portion of the Chernoff-Hoeffding bound. In that regard, one has the following result (see [2, Lemma 4.7.2]):

$$\mathbb{P}[\text{Bin}(n, p) \geq k] \geq \frac{1}{\sqrt{8k(1 - k/n)}} \cdot \exp(-nD(k/n||p)), \text{ for } k > np. \quad (3)$$

Briefly, this bound follows from the observation  $\mathbb{P}[\text{Bin}(n, p) \geq k] \geq \binom{n}{k} p^k (1 - p)^{n-k}$  combined with elaborate use of entropy estimates on binomial coefficients and appears to be among the best known lower bounds. Several lower bounds on the tail of binomial random variables can be found in [5, 8, 10, 15, 19], among many others. Let us remark that most lower bounds in the literature appear to be obtained using ideas from asymptotic analysis and/or rely on estimates of the normal approximation to the binomial distribution. In this article we provide yet another lower bound which results in a portion of the bound given by (2). We embark on a different approach that employs properties and estimates of tail conditional expectations.

Let us proceed by introducing our main result. Notice that the trivial estimates

$$p^k \leq \mathbb{P}[\text{Bin}(n, p) \geq k] \leq \binom{n}{k} p^k$$

imply that there exists a constant,  $C_{n,p,k}$ , that depends on  $n, p, k$  such that

$$\mathbb{P}[\text{Bin}(n, p) \geq k] = C_{n,p,k} \cdot p^k.$$

It turns out that  $C_{n,p,k}$  can be described in terms of tails conditional expectations of certain binomial random variables. More precisely, we will see in the next section that

$$C_{n,p,k} = \prod_{j=0}^{k-1} \frac{n-j}{\mathbb{E}[X_{n-j} | X_{n-j} \geq k-j]}, \text{ where } X_{n-j} \sim \text{Bin}(n-j, p)$$

and so the problem of obtaining lower bounds on the tail probability is reduced to the problem of obtaining upper bounds on tail conditional expectations of binomial random variables. The main novelty of our paper is that it appears to be the first report that investigates upper bounds on tail conditional expectations as a tool to obtain lower bounds on tail probabilities. More precisely, we obtain the following result. Here and later, given a real number  $x$ , we will denote by  $\lfloor x \rfloor$  the maximum integer that is less than or equal to  $x$ .

**Theorem 1.1.** *Let  $X_n \sim \text{Bin}(n, p)$  and fix a positive integer  $k$  such that  $np < k \leq n - 1$ . Set  $\ell = \lfloor \frac{k-np}{1-p} \rfloor < k$ . Then*

$$\mathbb{P}[X_n \geq k] \geq \frac{p^{2\ell}}{2} \cdot \binom{n}{\ell+1} / \binom{k}{\ell+1}.$$

In other words, Theorem 1.1 provides a lower bound on binomial tails that is equal to a certain portion of the upper bound given by (2). We prove Theorem 1.1 in the next section. Let us remark that comparisons between the bound provided by Theorem 1.1 and the bound given by (3) require quite tedious calculations. However, it is rather straightforward to put the computer to work and see that the bound of Theorem 1.1 sometimes performs better than the bound given by (3). We provide some pictorial comparisons between the two bounds in Section 3.

## 2 Proof of main result

In this section we prove our main result. The proof uses the following well-known result regarding the median of a binomial random variable.

**Theorem 2.1** (see [11]). *Let  $X \sim \text{Bin}(n, p)$ . Then  $\mathbb{P}[X \geq \lfloor np \rfloor] \geq 1/2$ .*

A basic ingredient in the proof of Theorem 1.1 is the following relation between the tails of a binomial random variable and its tail conditional expectation.

**Lemma 2.2.** *Fix a positive integer  $n$  and a real number  $p \in (0, 1)$ . For  $j = 1, \dots, n$  let  $X_j$  be a  $\text{Bin}(j, p)$  random variable. If  $k$  is a positive integer such that  $1 \leq k \leq n - 1$ . Then*

$$\mathbb{P}[X_n \geq k] = \frac{np}{\mathbb{E}[X_n | X_n \geq k]} \cdot \mathbb{P}[X_{n-1} \geq k - 1].$$

*Proof.* Since  $X_n = \sum_i B_i$  is a sum of  $n$  Bernoulli  $0/1$  random variables  $B_i$  of mean  $p$ , the linearity of expectation yields

$$\mathbb{E}[X_n | X_n \geq k] = \sum_{i=1}^n \mathbb{P}[B_i = 1 | X_n \geq k] = \frac{np \cdot \mathbb{P}[X_{n-1} \geq k - 1]}{\mathbb{P}[X_n \geq k]}$$

and the result follows.  $\square$

Notice that Lemma 2.2 can be iterated. This provides a way to express the tail probability of a binomial random variable in terms of tail conditional expectations.

**Corollary 2.3.** *Fix a positive integer  $n$  and a real number  $p \in (0, 1)$ . For every  $j \leq n$ , let  $X_j$  be a  $\text{Bin}(j, p)$  random variable. If  $\ell$  is any integer from the set  $\{0, 1, \dots, k-1\}$  then*

$$\mathbb{P}[X_n \geq k] = \mathbb{P}[X_{n-(\ell+1)} \geq k - (\ell + 1)] \cdot \prod_{j=0}^{\ell} \frac{p(n-j)}{\mathbb{E}[X_{n-j} | X_{n-j} \geq k-j]}.$$

*Proof.* Immediate upon iterating Lemma 2.2  $\ell + 1$  times.  $\square$

Hence lower bounds on the tail probability can be provided via upper bounds on the tail conditional expectations  $\mathbb{E}[X_{n-j} | X_{n-j} \geq k-j]$ , for  $j = 0, \dots, k-1$ . The next result provides a recursion for tail conditional expectations of integer valued random variables and is interesting on its own.

**Lemma 2.4.** *Let  $X$  be an integer valued random variable and fix a positive integer  $k$  such that  $\mathbb{P}[X \geq k] > 0$  and  $\mathbb{P}[X \geq k+1] > 0$ . Then*

$$\mathbb{E}[X | X \geq k] = k \cdot \frac{\mathbb{P}[X = k]}{\mathbb{P}[X \geq k]} + \mathbb{E}[X | X \geq k+1] \cdot \left(1 - \frac{\mathbb{P}[X = k]}{\mathbb{P}[X \geq k]}\right).$$

*Proof.* The result follows upon writing

$$\begin{aligned} \mathbb{E}[X | X \geq k] &= k \cdot \frac{\mathbb{P}[X = k]}{\mathbb{P}[X \geq k]} + \sum_{j=k+1}^n j \cdot \frac{\mathbb{P}[X = j]}{\mathbb{P}[X \geq k+1]} \cdot \frac{\mathbb{P}[X \geq k+1]}{\mathbb{P}[X \geq k]} \\ &= k \cdot \frac{\mathbb{P}[X = k]}{\mathbb{P}[X \geq k]} + \mathbb{E}[X | X \geq k+1] \cdot \left(1 - \frac{\mathbb{P}[X = k]}{\mathbb{P}[X \geq k]}\right). \end{aligned}$$

$\square$

We will also need the following result which provides estimates on ratios of binomial tails.

**Lemma 2.5.** *Let  $X_n \sim \text{Bin}(n, p)$ , for  $n = 1, 2, \dots$ . Then for every positive integer  $k$  such that  $np < k \leq n$  we have*

$$\mathbb{P}[X_n \geq k] \leq \frac{k(1-p)}{k-np} \cdot \mathbb{P}[X_n = k].$$

*In particular,*

$$\frac{\mathbb{P}[X_n \geq k+1]}{\mathbb{P}[X_n \geq k]} \leq \frac{p(n-k)}{k(1-p)}.$$

*Proof.* The first statement is a well known result. See, for example, Feller [7, p. 151] for a proof that employs comparison with a geometric series, or Diaconis et al. [5, Theorem 1] for a proof that employs the, so-called, Todhunter's formula. The second statement is equivalent to the first and follows upon observing that

$$\mathbb{P}[X_n \geq k+1] = \mathbb{P}[X_n \geq k] - \mathbb{P}[X_n = k].$$

For the sake of completeness, let us provide yet another proof of this result that employs Lemma 2.2. Notice that, if we apply Lemma 2.2 with  $n$  replaced by  $n+1$  and  $k$  replaced by  $k+1$ , we get

$$\begin{aligned} 0 \leq \mathbb{E}[\max\{0, X_{n+1} - (k+1)\}] &= \mathbb{E}[X_{n+1} - (k+1) | X_{n+1} \geq k+1] \cdot \mathbb{P}[X_{n+1} \geq k+1] \\ &= (n+1)p \cdot \mathbb{P}[X_n \geq k] - (k+1) \cdot \mathbb{P}[X_{n+1} \geq k+1] \\ &= p(n-k) \cdot \mathbb{P}[X_n \geq k] - (k+1)(1-p) \cdot \mathbb{P}[X_n \geq k+1]. \end{aligned}$$

Hence we have shown that

$$\frac{\mathbb{P}[X_n \geq k+1]}{\mathbb{P}[X_n \geq k]} \leq \frac{p(n-k)}{(k+1)(1-p)}$$

and the result follows.  $\square$

We can now proceed with an upper bound on the tail conditional expectation of a binomial random variable.

**Theorem 2.6.** *Let  $X \sim \text{Bin}(n, p)$  and fix positive integer  $k$  such that  $np < k \leq n$ . Then*

$$\mathbb{E}[X | X \geq k] \leq k + \frac{(n-k)p}{k - np + p}.$$

*Proof.* Clearly,  $\mathbb{E}[X | X \geq n] = n$ . Moreover, elementary calculations show that

$$\mathbb{E}[X | X \geq n-1] = n-1 + \frac{p}{(n-1)(1-p) + p} \leq n-1 + \frac{p}{n-1-np+p}$$

and so the result holds true for  $k \in \{n-1, n\}$ . For the remaining values, i.e.  $k \leq n-2$ , we prove the result by reverse induction on  $k$ . More precisely, we prove that the inequality holds true for  $k > np$  under the assumption that it holds true for  $k+1$ . The assumption that the inequality holds true for  $k+1$  implies that

$$\mathbb{E}[X | X \geq k+1] \leq k+1 + \frac{(n-k-1)p}{k+1-np+p}.$$

Now notice that the previous inequality combined with Lemma 2.4 yield

$$\begin{aligned}\mathbb{E}[X|X \geq k] &= k \cdot \frac{\mathbb{P}[X = k]}{\mathbb{P}[X \geq k]} + \mathbb{E}[X|X \geq k+1] \cdot \frac{\mathbb{P}[X \geq k+1]}{\mathbb{P}[X \geq k]} \\ &\leq k + \left(1 + \frac{(n-k-1)p}{k+1-np+p}\right) \cdot \frac{\mathbb{P}[X \geq k+1]}{\mathbb{P}[X \geq k]} \\ &\leq k + \left(1 + \frac{(n-k-1)p}{k+1-np+p}\right) \cdot \frac{p(n-k)}{k(1-p)},\end{aligned}$$

where the last estimate follows from Lemma 2.5. Now notice that the result will follow once we show

$$\left(1 + \frac{(n-k-1)p}{k+1-np+p}\right) \cdot \frac{1}{k(1-p)} \leq \frac{1}{k-np+p},$$

or, equivalently, once we show

$$\frac{k+1-kp}{k+1-np+p} \cdot \frac{1}{k(1-p)} \leq \frac{1}{k-np+p}.$$

Since we assume  $k < n-1$ , it is now easy to verify that the last inequality holds true and the proof is complete.  $\square$

We are now ready to prove our main result. Our proof combines the previous upper bound on tail conditional expectations with Corollary 2.3.

*Proof of Theorem 1.1.* Notice that  $\ell = \lfloor \frac{k-np}{1-p} \rfloor$  is the maximum integer  $j$  such that  $k-j \geq (n-j)p$ . From Corollary 2.3 we know that

$$\mathbb{P}[X_n \geq k] = \mathbb{P}[X_{n-(\ell+1)} \geq k - (\ell+1)] \cdot \prod_{j=0}^{\ell} \frac{p(n-j)}{\mathbb{E}[X_{n-j}|X_{n-j} \geq k-j]},$$

where  $X_{n-j} \sim \text{Bin}(n-j, p)$ , for  $j = 0, \dots, \ell$ . Since  $k - (\ell+1)$  is a positive integer such that  $k - (\ell+1) < (n - (\ell+1))p$ , Theorem 2.1 yields

$$\mathbb{P}[X_{n-(\ell+1)} \geq k - (\ell+1)] \geq \frac{1}{2}.$$

Now notice that Theorem 2.6 implies

$$\frac{p(n-j)}{\mathbb{E}[X_{n-j}|X_{n-j} \geq k-j]} \geq \frac{(n-j) \cdot p \cdot (k-j - (n-j)p) + (n-j)p^2}{(k-j) \cdot (k-j - (n-j)p) + (n-j)p}, \text{ for } j = 0, \dots, \ell.$$

Since  $k-j \geq (n-j)p$ , for  $j = 0, \dots, \ell$ , straightforward calculations show that

$$\frac{(n-j) \cdot p \cdot (k-j - (n-j)p) + (n-j)p^2}{(k-j) \cdot (k-j - (n-j)p) + (n-j)p} \geq \frac{n-j}{k-j} \cdot p^2.$$

Putting all the above together, we conclude that

$$\mathbb{P}[X_n \geq k] \geq \frac{1}{2} \cdot p^{2\ell} \prod_{j=0}^{\ell} \frac{n-j}{k-j}$$

and the result follows.  $\square$

### 3 Comparisons

In this section we perform pictorial comparisons between the bound given by Theorem 1.1 and the bound given by (3). In Figure 1 we fix the values of  $n$  and  $k$  and we plot the function  $f(p) := \frac{p^{2\ell}}{2} \binom{n}{\ell+1} / \binom{k}{\ell+1} - \frac{1}{\sqrt{8k(1-k/n)}} \cdot \exp(-nD(k/n||p))$ , for  $p \in (0, k/n)$ , where  $\ell$  is as in Theorem 1.1. For moderate values of  $p$  our bound is almost equal to the bound given by (3) while for large values of  $k$  and  $p$  our bound is sharper. In most cases, the two bounds are rather close to each other.

**Acknowledgements.** I am grateful to Henk Don for fruitful discussions and valuable comments.

### References

- [1] N. Alon, J. Spencer, *The Probabilistic Method*, 3rd Edition, J. Wiley and Sons, New York, 2008.
- [2] R.B. Ash, *Information Theory*, Dover, New York, (1990).
- [3] R.R. Bahadur, Some approximations to the binomial distribution function, *The Annals of Mathematical Statistics* **31**, no. 1, 43–54, (1960).
- [4] A. Carter, D. Pollard, Tusnády’s inequality revisited, *The Annals of Statistics* **32** no. 6, 2731–2741, (2004).
- [5] P. Diaconis, S. Zabell, Closed form summation for classical distributions: variations on a theme of de Moivre, *Statistical Science* **6** (3), 284–302, (1991).
- [6] W. Feller, On the normal approximation to the binomial distribution, *The Annals of Mathematical Statistics* **16**, no. 4, 319–329, (1945).
- [7] W. Feller, *An introduction to probability theory and its applications*, Volume 1, Wiley New York, (1957).

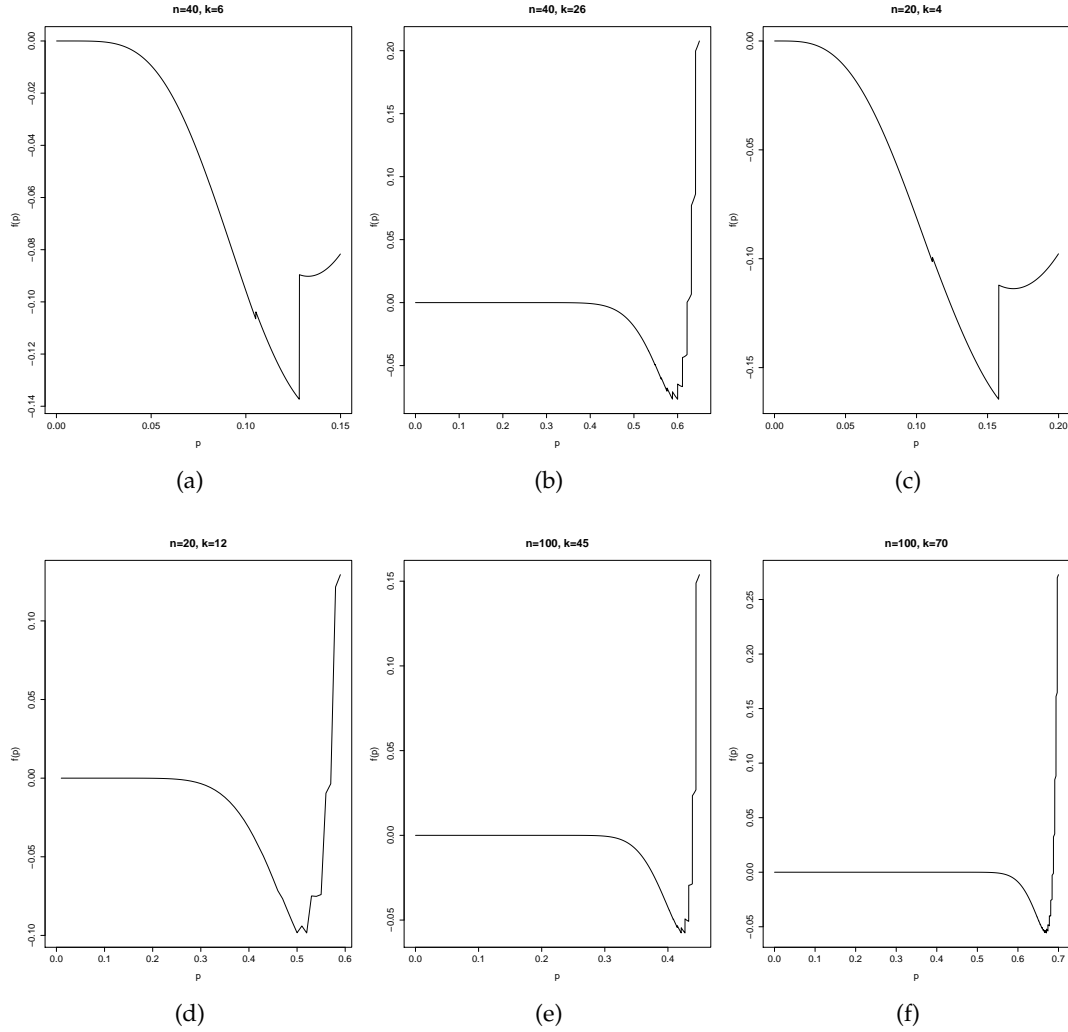


Figure 1: Comparison between the bound of Theorem 1.1 and the bound given by Eq.(3)

- [8] S. Greenberg, M. Mohri, Tight lower bound on the probability of a binomial exceeding its expectation, *Statistics and Probability Letters* **86**, 91–98, (2014).
- [9] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* **58**, p. 13–30, (1963).
- [10] E. Jeřábek, Dual weak pigeonhole principle, Boolean complexity, and derandomization, *Annals of Pure and Applied Logic* **129**, Issue 1–3, 1–37, (2004).



- [11] R. Kaas, J.M. Burhman, Mean, median, mode in the binomial distribution, *Statistica Neerlandica* **34** (1), 13–18, (1980).
- [12] J.E. Littlewood, On the probability in the tail of a binomial distribution, *Advances in Applied Probability* **1**, no. 1, 43–72, (1969).
- [13] P. Massart. Tusnady’s Lemma, 24 years later, *Annales de l’Institut Henri Poincare (B) Probability and Statistics* **38**, Issue 6, 991–1007, (2002).
- [14] B.D. McKay, On Littlewood’s estimate for the binomial distribution, *Advances in Applied Probability* **21**, no. 2, 475–478, (1989).
- [15] C. Pelekis, J. Ramon, A lower bound on the probability that a binomial random variable is exceeding its mean, *Statistics & Probability Letters*, (2016), DOI: 10.1016/j.spl.2016.08.016.
- [16] P. Rigollet, X. Tong, Neyman-Pearson classification, convexity and stochastic constraints, *Journal of Machine Learning Research* **12**, 2831–2855, (2011).
- [17] J.P. Schmidt, A. Siegel, A. Srinivasan, Chernoff-Hoeffding bounds for applications with limited independence, *SIAM Journal on Discrete Mathematics*, vol. 8, no. 2, p. 223–250, (1995).
- [18] E.V. Slud, Distribution inequalities for the binomial law, *The Annals of Probability* **5** (3): 404–421 (1977).
- [19] M. Telgarsky, Central binomial tail bounds, (preprint), arXiv:0911.2077.
- [20] A.M. Zubkov, A.A. Serov, A complete proof of universal inequalities for the distribution of the binomial law, *Theory Probab. Appl.* **57**, no. 3, 539–544, (2013).